DOCUMENT RESUME

ED 238 893 TM 820 470

AUTHOR Jones, Douglas H.; Ragosta, Marjorie
TITLE Predictive Validity of the SAT on Two Handicapped Groups: The Deaf and the Learning Disabled.
INSTITUTION Educational Testing Service, Princeton, NJ. Program Statistics Research Project.
REPORT NO ETS-PSR-TR-82-27; ETS-RR-82-9
PUB DATE Jan 82
NOTE 66p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *College Entrance Examinations; *Deafness; *Grade Point Average; *Learning Disabilities; Performance Factors; Predictive Measurement; *Predictive Validity; *Scores; Test Validity
IDENTIFIERS *Scholastic Aptitude Test

ABSTRACT

The purpose of the study was to evaluate the predictive validity of the Scholastic Aptitude Test (SAT) for deaf students at one institution and learning disabled (LD) students at another. SAT scores and measures of high school and college performance were obtained at each institution for both handicapped and nonhandicapped students. Additional descriptive information was obtained for the handicapped groups. The SAT scores of deaf students (when available) were significantly lower than the verbal and mathematical scores of the hearing students and tended to underpredict deaf students' college performance. The combination of high school grade point average and SAT scores predicted performance equivalently for deaf and hearing students. For the learning disabled, the SAT verbal score and the high school and college performance measures were all significantly lower than those for the non-LD students. The SAT validity coefficients for the two groups were not statistically significantly different. The combination of high school rank and SAT scores produced a significant lower validity coefficient for LD students than for non-LD students. Since these were preliminary studies, no implications are being drawn pending further research. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED230893

Collective Validity of Utilities of Correct Decisions: The Decision-Theoretic Evaluation of Selection Procedures

Douglas N. Jones
and
Melvin R. Novick

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Wildemmiller

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

PROGRAM STATISTICS RESEARCH

TECHNICAL REPORT NO. 80-7

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY 08541

Predictive Validity of the SAT on Two Handicapped Groups: The Deaf and the Learning Disabled

Douglas H. Jones

and

Marjorie Ragosta

Program Statistics Research
Technical Report No. 82-27

ETS Research Report
RR-82-9

January 1982

with the assistance of

Jill A. Callahan, Linda J. DeLauro and Norma Norris

Copyright © 1982 by Educational Testing Service. All rights reserved.

The Program Statistics Research Technical Report Series is designed to make the working papers of the Research Statistics Group at Educational Testing Service generally available. The series consists of reports by the statisticians in the Research Statistics Group as well as their external and visiting statistical consultants.

Reproduction of any portion of a Program Statistics Research Technical Report requires the written consent of the author(s).

Abstract

Because of the ambiguity of Section 504 of the Rehabilitation Act of 1973, with regard to the need for predictive-validity studies of admissions tests for handicapped persons, a preliminary study was undertaken. The purpose of the study was to evaluate the predictive validity of the SAT for deaf students at one institution and learning disabled (LD) students at another. SAT scores and measures of high school and college performance were obtained at each institution for both handicapped and nonhandicapped students. Additional descriptive information was obtained for the handicapped groups. The SAT scores of deaf students (when available) were significantly lower than the verbal and mathematical scores of the hearing students and tended to under predict deaf students' college performance. The combination of high school grade point average and SAT scores predicted performance equivalently for deaf and hearing students. For the learning disabled, the SAT verbal score and the high school and college performance measures were all significantly lower than those for the non-LD students. The SAT validity coefficients for the two groups were not statistically significantly different. The combination of high school rank and SAT scores produced a significant lower validity coefficient for LD students than for non-LD students. Since these were preliminary studies, no implications are being drawn pending further research.

Acknowledgments

The very special help of two people and two institutions must be acknowledged at the outset. To Ray Jones, Director of the National Center on Deafness at California State University, Northridge (CSUN), we extend our thanks for his efforts to help us understand the deaf and for being the first to volunteer to help on such a project. Without his support, this project might not have been undertaken. To Gertrude Webb, Director of the Program of Assistance in Learning at Curry College, we owe thanks for her efforts to help us understand the learning disabled and for volunteering her support.

To all of you at CSUN and Curry College who helped us obtain the best data we could hope for, many thanks. We promise you we will continue to try to shed as much light as possible on the problems of handicapped students and standardized tests.

Table of Contents

		Page
1.	Chapter 1: Introduction	1
2.	Chapter 2: SAT Scores and the Deaf	4
3.	Chapter 3: SAT Scores and the Learning Disabled. . .	23
4.	Chapter 4: Next Steps.	36
	References.	40
	Appendix A	
	Appendix B	

Chapter 1

INTRODUCTION

The Rehabilitation Act of 1973 affected profound changes in the education of handicapped persons. One aspect of Section 504 concerned admissions and recruitment of handicapped persons to post-secondary education. Section 84.42 stated, among other things, that an educational institution receiving federal assistance:

> (2) may not make use of any test or criterion for admission that has a disproportionate, adverse effect on handicapped persons or any class of handicapped persons unless (i) the test or criterion, as used by the recipient, has been validated as a predictor of success in the education program or activity in question and (ii) alternate tests or criteria that have a less disproportionate, adverse effect are not shown by the Director to be available;
>
> (3) shall assure itself that (i) admissions tests are selected and administered so as best to ensure that, when a test is administered to an applicant who has a handicap that impairs sensory, manual or speaking skills, the test results accurately reflect the applicant's aptitude or achievement level or whatever other factor the test purports to measure, rather than reflecting the applicant's impaired sensory, manual, or speaking skills (except where those skills are the factors that the test purports to measure);
>
> (4) except as provided (in paragraph (c) of this section) may not make preadmission inquiry as to whether an applicant for admission is a handicapped person.

Several problems arose from these regulations. The wording of the regulation implementing Section 504 was ambiguous regarding the group of people on which validity studies should be performed. Were special studies of validity required for the handicapped? Were the regulations requiring tests of aptitude or achievement independent of sensory, manual or verbal handicaps? Would testing companies be allowed to flag scores of tests taken by handicapped persons in special test administrations if, by flagging, they were identifying students as handicapped?

A Panel of the Testing of Handicapped People was proposed by the National Academy of Sciences and funded by the Office of Civil Rights. A major product of the Panel was a report to be released in the fall of 1981 and to be comprised of three main sections: a full discussion of the issues, a detailed research agenda and recommendations for modifications in policies consistent with the current state of the art. In testimony presented to the Panel on March 14, 1980, Educational Testing Service (ETS) proposed several categories of research to be undertaken on as rapid a schedule as possible:

> One of the immediate research needs in the testing of the handicapped is that of determining the extent to which the tests are predictive of academic performance in undergraduate or graduate and professional schools. Also needed is a comparative study of the predictive power of the tests for handicapped and nonhandicapped students. A serious difficulty in carrying out such studies, however, is the paucity of students who are homogeneous with respect to certain relevant characteristics, a condition necessary for the conduct of meaningful validity studies.
>
> What is needed are groups of students of sufficient number who share the same handicap and are enrolled in the same institution and in the same academic year, a set of restrictions necessary for the assurance of reasonably consistent grading standards.

In his report, summarizing population validity studies, Breland (1979) noted that validity studies, explicitly using data from handicapped populations, do not exist.

As part of ETS' continuing concern for handicapped students and the validity of its admissions tests, the current study was undertaken. Data bases were sought from colleges where large numbers of handicapped students participate in course work with nonhandicapped students.

Two sources were found: a university which provided support services to mainstreamed deaf students and a college where learning-disabled students attended regular classes but in addition participated in a program of assistance in learning. Chapter 2 contains information on deaf students and the SAT. Chapter 3 contains information on learning-disabled students and the SAT.

Chapter 2

SAT SCORES AND THE DEAF

Identification numbers and information on handicapping conditions for approximately 500 deaf students were supplied to ETS by a large university. Information included the classification of students as deaf or hard-of-hearing and the age of onset of deafness. This information was used in two ways: to provide the basis for a description of the sample of deaf students and to produce a finder's deck to search the university files for first year average, high school grade point average and SAT scores.

Description of the Deaf Students

Cause of Deafness

The self-reported causes of deafness in students attending the university are reported in Table II.1. Almost half of the students listed the cause as unknown. The most frequently reported cause was infection or illness. Thirteen percent--68 students--defined their deafness as congenital or hereditary.

Table II.1

Causes of Deafness in the Student Sample

Congenital, hereditary	68 students
Maternal rubella	22 students
Birth incidents	21 students
Infections, illness	116 students
Accident	11 students
Nerve damage (undefined)	40 students
Unknown	238 students
Total	516 students

Age of Onset of Deafness

A frequency distribution of students who had data on the age of onset of deafness and reported SAT scores is given in Table II.2. For the verbal SAT scores, clustering near the bottom of the score distribution is evident, especially in the group listing birth as the age of onset. Almost half of the students in that category had scores in the 200s--the lowest range of scores. The nine students whose age of onset was listed as unknown had a distribution of verbal SAT scores which most resembled the birth group although they could not be classified.

Table II.2

The Frequency Distribution of Verbal (V) and Mathematical (M) SAT Scores Based on the Age of Onset of Deafness (N = 97 Students)

SAT Scores	Age of Onset of Deafness							
	Birth		Preschool		5 or After		Unknown	
	V	M	V	M	V	M	V	M
200s	23	5	5	3	4	4	5	4
300s	18	21	7	5	3	6	1	1
400s	5	17	4	5	1	4	1	3
500s	4	7	3	4	6	4	2	0
600s	0	0	1	3	3	0	0	1
700s	0	0	0	0	1	0	0	0
Ns	50		20		18		9	

The shift of verbal scores upward with the age of onset is made even clearer in Table II.3. Students were classified as deaf or hard-of-hearing by the university staff responsible for support services. It is important to note that even those students classified as hard-of-hearing were so severely

impaired that ordinary speech was largely unintelligible There were 82 undergraduate students for whom there existed SAT data, a deaf or hard-of-hearing classification, and information on the age of onset of deafness. The data show smaller differences in verbal SAT scores between the deaf and hard-of-hearing than the differences associated with the age of onset of deafness. A mean score of 319 was obtained for the 46 students who were deaf or hard-of-hearing from birth. The scores improved markedly with later ages of onset.

Table II.3

SAT Verbal and Mathematical Scores for Deaf and Hard-of-Hearing Students Classified by Age of Onset of Deafness

	Deaf			Hard-of-Hearing			Total		
Verbal	(N)	$\bar{X}$	SD	(N)	$\bar{X}$	SD	N	$\bar{X}$	SD
Age of Onset									
Birth	(25)	316	85	(21)	323	114	(46)	319	100
Preschool	(8)	361	103	(10)	400	117	(18)	383	109
5 & Over	(8)	443	156	(10)	484	178	(18)	466	165
V Total	(41)	349	114	(41)	381	147	(82)	365	131
Mathematical									
Age of Onset									
Birth	(25)	394	71	(21)	394	93	(46)	394	81
Preschool	(8)	417	150	(10)	469	114	(18)	446	130
5 & Over	(8)	377	121	(10)	413	104	(18)	397	110
M Total	(41)	395	99	(41)	417	103	(82)	406	100

The Validity Study

As previously mentioned, student identification numbers provided by the university were used to produce a finder's deck of computer cards. The finder's deck contained not only the ID numbers for deaf students but also a randomly generated set of ID numbers for hearing students who entered the university during the same semesters as the deaf students. The finder's deck was used to search the university files to obtain the following data for each ID number: high school grade point average (HSGPA), the first-year college grade point average (FYA) and the SAT verbal (V) and mathematical (M) scores. This information provided the data for this validity study.

Missing Data

The number of student records available for the differential validity study is 141 hearing students and 60 deaf students. Of the other 1,091 hearing students and 540 deaf student cases, 510 hearing and 244 deaf nontransfer students are missing values on one or more variables and the remainder are transfer students who have incompatible FYA's. Table II.4 shows a breakdown of students into groups with specific missing and available data.

Table II.4

Breakdown of Cases into Categories of Missing Data

Variables				Number of Cases							
				Hearing				Deaf			
Transfer (?)	SAT	HSGPA	FYA	Counts	Percentages+			Counts	Percentages+		
					A	B	C		A	B	C
No	1	1	1	141	11	22	53	60	10	15	32
No	1	1	0	49	04	08		29	05	07	
No	1	0	1	7	00	01	03	12	02	03	06
No	1	0	0	24	02	04		5	00	01	
No	0	1	1	4	00	01	02	13	02	03	07
No	0	1	0	5	00	01		8	01	02	
No	0	0	1	112	09	17	42	101	17	25	54
No	0	0	0	309	25	48		176	29	44	
Yes	1	1	1	58	05			19	03		
Yes	1	1	0	16	01			1	00		
Yes	1	0	1	5	00			3	00		
Yes	1	0	0	0	00			1	00		
Yes	0	1	1	0	00			4	00		
Yes	0	1	0	2	00			3	00		
Yes	0	0	1	389	32			141	24		
Yes	0	0	0	111	09			24	04		

1 = Not Missing 0 = Missing

+Denominators for determining proportions in columns A, B, and C are as follows: A, total number of both transfer and nontransfer students; B, total number of nontransfer students; C, total number of nontransfer students with FYA.

Since there is a considerable amount of missing data, there is concern that the results based on the complete data could be radically different if the missing data were available. Concern would be lessened if the incomplete data became missing by a random mechanism. Another condition for valid inferences based on incomplete information is for the probability of a variable having a missing value not to depend on the value of the variable (Rubin, 1976). There are two mechanisms by which the score may be missing because of the value or potential value of the score. Students having a 3.2 HSGPA or larger, may gain admittance to CSUN without SAT scores. In addition, deaf students may gain admittance through special procedures without regard to admissions credentials. Perhaps, some hearing students may gain admittance by similar exemptions from the standard admissions procedures, some of which may account for 112 hearing students without credentials but with FYA. The percentage of hearing and deaf nontransfer students with complete data are, respectively, 11% and 10% relative to the total data base. These appear to be small percentages; however, the percentages increase to 53% and 32% when the data base is limited to nontransfer students with FYA's.

The set of incomplete data, that is missing the FYA score, is compared with complete data in Table II.5 on the values of the means, standard deviations and correlation matrices of each variable. The means of SAT-V, SAT-M, and HSGPA are similar, but the correlation matrices are not. Under the assumption that the incomplete data were generated randomly from the same source as the complete data, the hearing group correlations between SAT-V and SAT-M and between SAT-V and HSGPA are lower than expected; further, the deaf group correlation between SAT-V and SAT-M is higher than expected, and the correlation between SAT-M and HSGPA is lower than expected. Tests of significance, one each for the deaf and

Table II.5

Comparison of SAT Verbal and Mathematical Scores and High School Grade Point Averages for Hearing and Deaf Students with First Year Averages (Table A) and without (Table B), Including Correlations among the Variables

Table A

Complete Data

	Mean (S.D.)		Correlation		
	Hearing	Deaf	Hearing (140)		Deaf (60)
			SAT-V	SAT-M	HSGPA
FYA	2.65(0.73)	2.49(0.61)	0.3814/0.1380	0.3172/0.4059	0.6315/0.5156
SAT-V	449 (98)	347 (109)		0.5985/0.4068	0.3087/-0.0174
SAT-M	489 (107)	409 (100)			0.2833/0.3566
HSGPA	3.20 (0.42)	3.20 (0.39)			

Table B

Incomplete Data (Missing FYA)

	Mean (S.D.)		Correlation	
	Hearing	Deaf	Hearing (119)	Deaf (60)
			SAT-M	HSGPA
SAT-V	462 (86)	365 (109)	0.5053/0.6394	0.1990/0.0570
SAT-M	469 (99)	416 (100)		0.2270/0.1605
HSGPA	3.12 (0.37)	3.05 (0.41)		

hearing, of equal covariance matrices are marginally significant with P-values approximately 0.10 (the test is described on page 252 of Morrison, 1976).

Also compared are the mean FYA of the complete data and the mean FYA of the set of incomplete data missing the SAT and/or HSGPA. Table II.6 shows the values of the means and t-statistics. The incomplete data FYA is higher in both the deaf and hearing groups. An explanation for this phenomenon is unavailable from the data.

Table II.6

College FYA for Hearing and Deaf Students with Complete and Incomplete Data on SAT Scores and HSGPA, including Test of Differences between Complete and Missing Data

	Complete Data			Incomplete Data (Missing SAT and/or HSGPA)			Tests of Differences		
Hearing	N	MEAN	SD(N-1)	N	MEAN	SD(N-1)	POOL SD	T	DF	P
FYA	140	2.658	0.726	112	3.362	0.649	0.693	-8.01	250	0.000
Deaf										
FYA	60	2.485	0.613	101	3.161	0.480	0.533	-7.77	159	0.000

This section is concluded with the reminder that the results on the non-comparability of the complete and incomplete data do not imply that the information on differential validity from the complete data is in error.

Results

Results on the differential performance between the deaf and hearing will be presented first. These will be followed by presentation of the validity coefficient for each group. And these in turn will be followed by presentation of separate regression equations, predicting FYA, for the deaf and hearing groups. This will comprise the analysis of differential prediction and validity.

Table II.7 shows the summary statistics on the SAT and HSGPA for both groups. Deaf students have significantly lower SAT scores than hearing students do, but both groups have equivalent distributions of HSGPA. The distributions of SAT-V and SAT-M scores for the deaf group are approximately 100 and 68 points lower, respectively, than the hearing group; and the multivariate Hottelling's T^2 test of differences on SAT-V, SAT-M, and HSGPA yields an F statistic of 22.2 which has a P-value less than .005 (3,363 degrees of freedom).

A discriminant analysis (see Table A, Appendix A) was performed using the SAT and HSGPA to see how well these variables can discriminate between the deaf and hearing. It was found that the discriminating power was very poor, yielding a 32% classification error rate. This implies that although the distribution of scores for the two groups is significantly different, there are considerable numbers of deaf students scoring higher than hearing students. Scatterplots of the data are also included as Figure B, Appendix A.

Table II.7

Means, Standard Deviations, and t-statistics of the SAT-V, SAT-M and HSGPA Variables for the Deaf and Hearing

Variable	Means (S.D.) Hearing (N = 259)	Deaf (N = 108)	Difference (S.D.)	t-statistic
SAT-V	455.10 (92.56)	355.28 (112.88)	99.81 (98.91)	8.84
SAT-M	480.46 (100.33)	413.50 (100.96)	67.91 (102.65)	5.79
HSGPA	3.16 (0.59)	3.13 (0.62)	0.03 (0.60)	0.44

Table II.8 presents the validity coefficient of the various prediction systems: the HSGPA alone, the SAT composite (two predictors SAT-V and SAT-M), and the HSGPA together with both SAT scores. The values of the HSGPA and SAT-M

validity coefficients for the two groups are not statistically significantly different; however, the SAT-V validity coefficient is significantly lower for the deaf compared to the hearing. It appears that the SAT-V score is not as useful in predicting college performance for the deaf as it is for the hearing. Although the SAT-V validity coefficient for the deaf is small, 0.14, it would be premature to conclude that the true validity coefficient is 0 since the data set is so small (N = 60).

The square of the multiple correlation coefficient, R-square, for each predictor system is the percent of variance in FYA explained by that predictor system. The increment in R-square due to adding more predictors is the relative percent explained in the variance of FYA. For the deaf, the relative percent of variance (in FYA) explained by adding SAT to HSGPA and by adding

Table II.8

Validity Coefficients of Various Prediction Systems for the Deaf and Hearing Students

Prediction System	Validity Coefficient R (R^2) Hearing (N = 140)	Deaf (N = 60)	P-Value of Test of Difference between Deaf and Hearing Coefficients
1. HSGPA	0.63 (0.40)	0.52 (0.27)	0.250
2. Composite Test Scores	0.47 (0.22)	0.43 (0.19)	0.390
SAT-V	0.38 (0.14)	0.14 (0.02)	0.059
SAT-M	0.32 (0.10)	0.41 (0.17)	0.337
3. HSGPA and Tests	0.66 (0.44)	0.58 (0.34)	0.215

HSGPA to SAT is 7% and 15%. The former value is marginally significant ($F = 2.99$, d.f. = 2,56; $.05 < P < .10$) and the latter value is highly significant ($F = 12.45$, d.f. = 1,56; $P << .005$). For the deaf, the percent of variance contributed by the SAT is almost entirely accounted for by the SAT-M score. For the hearing, the relative percent of variance (in FYA) explained by adding SAT to HSGPA and adding HSGPA to SAT is 3% and 22%. The relative strengths almost duplicate those for the deaf: for the SAT, $F = 3.77$, d.f. = 2,137, $.05 < P < .10$; and for HSGPA, $F = 25.13$, d.f. = 1,137, $P << .005$. The fact that the increment in R-square due to adding SAT to HSGPA is greater for the deaf than the hearing is for the most part due to the poorer predictive value of the HSGPA for the deaf.

Attention is now turned to the important issue of whether use of prediction systems based on the hearing group's data has a negative impact on the deaf. This question will be answered in two ways. The first approach will be to estimate and compare two separate prediction formulas, one each for the deaf and hearing groups. These formulas will be based on the combined data across groups, rather than each formula being obtained from separate data sets (see Gulliksen and Wilk, 1950, for a complete description). The reason for this approach is that the parameters are estimated with greater accuracy based on combined data than can be obtained from separate group data. The second approach will be to analyze deaf students' predicted FYA's obtained from a formula based only on hearing students' data. The former approach is discussed first.

Until now the analyses have made use of only three variables to predict FYA: HSGPA, SAT-V, and SAT-M. Separate prediction formulas for the deaf and hearing are estimated by the introduction of a fourth predictor, denoted by

HAND, indicating whether or not an individual has a disability. In addition, HAND is multiplied together with each of the other predictors to produce three more variables in order to describe interactions between scores and the handicap. Separate prediction formulas are then obtained for an individual by using his/her value of the predictor HAND. Significant differences in prediction formulas between the deaf and hearing are determined by evaluating the contribution to R-square that the additional predictors (HAND and interactions) make.

Table II.9 shows the R-square values for each of three prediction systems, with and without the variable HAND. Values of the regression weights and related statistics associated with each of the three prediction systems are given in Table C, and scatterplots are included as Figure D, Appendix A. Each system is discussed separately.

Table II.9

R-square Value of the Three Prediction Systems With and Without the Handicap Indicator

	Prediction System	R-square
1.	HSGPA	0.36
	HSGPA and HAND	0.38
2.	Composite Test Score	0.14
	SAT-V	0.11
	SAT-M	0.13
	SAT Composite Test Scores and HAND	0.17
3.	HSGPA with SAT Composite Test Scores	0.41
	HSGPA, SAT Composite Test Scores and HAND	0.42

The HSGPA. The HSGPA prediction system does not operate equivalently for the deaf and hearing. The contribution to R-square due to the addition of the variable HAND is 0.02 yielding an F-statistic equal to 2.97 with 2 and 196 degrees of freedom ($0.05 < P < 0.10$).

By use of the prediction formula based on the hearing group's data, a deaf student would be predicted to have a higher FYA than his/her own equation would indicate. (See Appendix A, Table D.) A deaf student with HSGPA equal to 3.13, the deaf group mean, would be predicted to obtain an FYA equal to 2.58 based on the hearing group's data; whereas, the same student would on the average obtain an FYA equal to 2.43. Even though the overall test of different regressions between the two groups is statistically significant, at the deaf group's mean the difference in predicted FYA, 0.15, is not significant; since the standard error of estimate is approximately 0.20, this means that the practical differences in predicted FYA would be observed only for outlying values of HSGPA.

SAT-V and SAT-M. The SAT composite prediction system, like the HSGPA system, also does not operate equivalently for the deaf and hearing students. In Table II.9, the increment in R-square, 0.03, due to the addition of the variable HAND yields an F-statistic equal to 4.82 with 1 and 194 degrees of freedom ($0.025 < P < 0.05$). The net effect of predicting deaf students' FYA based on the hearing group's data is either to over or under predict depending on the particular score combination of SAT-V and SAT-M.

A mode of looking at this result is to determine whether the operational inequivalency of the SAT produces significant differences in predicted FYA's obtained from "hearing" and "deaf" formulas. If significantly higher predictions exist for the deaf with the use of the deaf prediction system, then a negative impact would occur from use of the "hearing" prediction system.

Table II.10 shows the differences between deaf and hearing formulas in predicted FYA evaluated at selected values of SAT-V and SAT-M. Positive values show under predictions and negative values show over predictions. The differences corresponding to typical SAT scores for the deaf are outlined by the dashed border in Table II.10. A difference significant at the 5% level is denoted by an asterisk. It is seen that use of the "hearing" prediction formula amounts to under predicting the deaf by approximately one half of a grade point for deaf students having SAT-V scores 300 or lower and SAT-M scores 400 through 600. Since those scores are close to the means presented in Table 11.3 for students deaf from birth, the under prediction of deaf performance by the use of the SAT scores alone is a matter for concern.

Why does the SAT under predict deaf students' performance? One hypothesis is that, under ordinary circumstances, it would not. Most students scoring at the 300 level or below on the verbal portion of the SAT would have a difficult time in college. For the deaf students in this study a wide range of support services is provided. Without those support services the FYA's of deaf students might have been better predicted by the SAT. Where colleges and universities provide support services for the deaf, the low SAT-V scores of deaf students may not be indicative of poor college performance.

The HSGPA and the SAT. In Table II.9, the prediction system combining the SAT and HSGPA operate equivalently for the deaf and hearing students. The contribution to R-square due to the addition of the variable HAND is 0.01, yielding an insignificant F-statistic equal to 0.69 with 4 and 192 degrees of freedom ($0.75 < P$). The SAT-V score under predicts and HSGPA over predicts FYA for deaf students, thus cancelling the individual effects. These findings are further confirmed in Table II.11 which shows no significant differences between the "hearing" and "deaf" formulas using the SAT plus the HSGPA prediction system.

A suitable fit of the linear model used in this last prediction system is indicated by the residual plots of Figure E in Appendix A. In addition to reasonable fit, residual (observed minus predicted) plots for the deaf students using the hearing students' prediction system reveal no systematic under or over prediction for the entire range of predicted FYA values. This further indicates that there is no inequivalency in the SAT plus HSGPA prediction system between hearing and deaf students. Separate plots for students deaf before or after 18 months show no peculiar patterns occur that may be due to age of onset of deafness.

Table II.10

Difference between Predicted FYA (Deaf Minus Hearing)
for Selected Values of SAT-V and SAT-M
Using the SAT Prediction System

Difference (S.E.)

SAT-V	200	300	400	500	600	700	800
800	-1.155(0.631)	-0.992(0.552)	-0.829(0.486)	-0.666(0.440)	-0.503(0.418)	-0.340(0.426)	-0.177(0.462)
700	-0.934(0.533)	-0.771(0.450)	-0.608(0.381)	-0.445(0.337)	-0.282(0.327)	-0.119(0.353)	0.044(0.410)
600	-0.713(0.440)	-0.550(0.351)	-0.387(0.279)	-0.224(0.240)	-0.061(0.250)	0.102(0.304)	0.265(0.383)
500	-0.492(0.356)	-0.329(0.261)	-0.166(0.185)	-0.003(0.160)	0.160(0.205)	0.323(0.288)	0.486(0.387)
400	-0.271(0.289)	-0.108(0.191)	0.055(0.120)	0.218(0.131)	0.381(0.212)	0.544(0.313)	0.707(0.419)
300	-0.050(0.254)	0.113(0.169)	0.276(0.134)	0.439(0.180)*	0.602(0.268)*	0.765(0.369)	0.928(0.475)
200	0.171(0.262)	0.334(0.211)	0.497(0.214)*	0.660(0.268)*	0.823(0.350)*	0.986(0.446)	1.149(0.548)

SAT-M

*The actual prediction system used for the deaf in constructing Table II.10 puts zero weight on the SAT-V score. The reason for this is that the least squares estimate of the SAT-V regression weight is -0.0018, which is not significantly different from zero and negative weights, which penalize students with higher SAT-V scores than others, would in all probability never be used in practice.

Table II.11

Difference between Predicted FYA (Deaf Minus Hearing) for Selected Values of SAT-V and SAT-M, Holding HSGPA = 3.0, Using the SAT Plus HSGPA Prediction Systems

Difference (S.E.)

SAT-V	SAT-M 200	300	400	500	600	700	800
800	-0.635(0.540)	-0.529(0.470)	-0.424(0.413)	-0.318(0.375)	-0.212(0.360)	-0.107(0.373)	-0.001(0.411)
700	-0.534(0.456)	-0.428(0.383)	-0.322(0.324)	-0.217(0.288)	-0.111(0.286)	-0.005(0.317)	0.100(0.372)
600	-0.432(0.376)	-0.327(0.298)	-0.221(0.237)	-0.115(0.209)	-0.010(0.226)	0.096(0.280)	0.202(0.355)
500	-0.331(0.304)	-0.225(0.220)	-0.120(0.158)	-0.014(0.146)	0.092(0.194)	0.197(0.272)	0.303(0.361)
400	-0.230(0.246)	-0.124(0.161)	-0.018(0.105)	0.087(0.128)	0.193(0.204)	0.298(0.295)	0.404(0.390)
300	-0.128(0.215)	-0.023(0.143)	0.083(0.121)	0.189(0.170)	0.294(0.251)	0.400(0.343)	0.505(0.438)
200	-0.027(0.222)	0.079(0.180)	0.185(0.188)	0.290(0.242)	0.395(0.319)	0.501(0.406)	0.607(0.498)

Discussion

The analysis of the data reveals that the HSGPA over predicts the FYA for most deaf students, and the opposite holds for the SAT prediction system. This over and under prediction may explain why the combined SAT and HSGPA predicts FYA equivalently for both the deaf and hearing students at the university.

For the deaf students, the SAT-V score does not measure verbal ability in the usual way. A low verbal score for nonhandicapped students usually indicates a lessened ability to perform well academically. However, the deaf students' data indicate that, for given SAT-M scores, high SAT-V scores are not needed to perform well.

The HSGPA's for the deaf and hearing are not comparable. One might speculate that many deaf students go to special schools for the deaf where curriculum and grading practices may differ from public and private high school. The data show that the HSGPA is overly optimistic for the deaf in predicting college FYA. This is true despite the fact that the university has a program of support services that is outstanding: interpreters and note-takers are routinely provided for deaf students attending any college course.

Although the data reveal that the deaf students' verbal scores are on the average 100 points lower than the hearing students' scores and have no correlation with HSGPA and low correlation with FYA, the SAT-V score may in fact measure the deaf students' ability to comprehend test materials in English. Low SAT-V scores may be an indication that other test scores, e.g., SAT-M, may also be depressed. However, the analysis of the SAT prediction system revealed that extra credit had to be given to the SAT-M score of deaf students in order to correctly predict FYA. The correlation between SAT-V and SAT-M for the deaf is nearly as high as it is for the hearing. Under the assumption that FYA

measures college performance equivalently for deaf and hearing students, these data suggest that the diminished SAT-M score does not measure the deaf students' full college potential. A possible explanation could be that the deaf students' low verbal abilities restrict their attempts even to demonstrate math skills on the SAT-M.

We have noted differences in the ability of the HSGPA alone, or SAT alone, to predict FYA's of deaf and hearing students. Although HSGPA, SAT-V and SAT-M are valid predictors of FYA for both groups, they operate differently and the information they contain should be interpreted differently with regard to deaf students. It is to be hoped that future admissions practices will incorporate knowledge on matters involving the deaf, since these data indicate that policies made for hearing students are inappropriate for decisions concerning deaf individuals at this university.

Chapter 3

SAT SCORES AND THE LEARNING DISABLED

The Validity Study

A small private college with a special program for learning-disabled students provided data for the following study. For the freshman class of 1979, learning-disabled (LD) and non-LD students were identified and the following data obtained: high school rank (HSR), SAT scores (SAT-V and SAT-M), and first semester grade point averages. In addition, for the learning-disabled students subtest scores on the Wechsler Adult Intelligence Scale (WAIS) were obtained. In the first section of this chapter a description of the learning-disabled students will be presented. The second section will contain a validity study of the SAT for the students in this data base.

A Description of the Learning-Disabled

What are the characteristics of the learning-disabled students at this small college? Their program director describes them as follows:

> Our population includes a football hero and a hockey star but also several boys who need help in body coordination. The boys outnumber the girls.
>
> ...Orally most communicate well making their points effectively although a few have expressive language problems and appear on first meeting to be very quiet. Overall they are curious, interested, intelligent and eager to participate in oral discussions. Many exhibit fine motor problems, difficulties in their visual and auditory integration systems.
>
> As a result of these handicapping conditions, though their oral performance is adequate, their written performance lacks any order. Sentences are fragmented--ideas are not sequenced--handwriting is illegible--spelling bizarre. Punctuation and capitalization virtually are nonexistent, incorrect at best. As with the written word, most are several grade levels behind in their reading ability. Both speed and comprehension are generally depressed.

> Even as we generalize about the characteristics of the LD population, we hasten to explain that no LD student has all the handicapping conditions described. One may have physical prowess and need help in building auditory skills so that he can better take in his professors' remarks. His neighbor, however, may have fine motor coordination and visual memory difficulties which cause him not to be able to write nor to easily remember much of what he has read. Still a third may be in trouble getting her ideas in order so that the teacher can follow them, while yet a fourth needs getting his ideas expressed on paper.

The WAIS subtest scores for 58 students in this study are given in Table III.1 which compares them with 17 normal and 57 LD college students recently described by Cardoni et al. (1981) in the Journal of Learning Disabilities.

In the Cardoni study, LD students scored lower than non-LD students on all subtests except Block Design and Object Assembly. The ETS LD group scored closer to the Cardoni LD group than to the normal group on all subtests except, again, Block Design and Object Assembly. Both LD groups reported Full Scale IQs of 108.

Table III.1

Means and Standard Deviations of WAIS Scores

	Cardoni et al. Study				ETS LD Study	
	Normal		Learning Disabled		Learning Disabled	
	M	SD	M	SD	M	SD
Information	11.6	1.4	9.6	2.0	9.7	2.1
Comprehension	12.9	2.9	12.5	3.0	12.4	2.7
Arithmetic	10.9	2.7	9.5	3.0	10.1	2.8
Similarities	13.8	2.4	12.3	2.5	12.0	2.3
Vocabulary	11.8	1.6	10.7	2.0	10.2	1.7
Digit Span	12.7	3.1	9.2	2.9	9.9	2.6
Digit Symbol	12.1	2.2	10.0	2.4	10.9	3.0
Picture Completion	12.1	2.3	11.5	1.9	11.4	2.8
Picture Arrangement	11.8	2.4	10.9	1.8	10.7	2.0
Block Design	12.1	2.4	12.2	3.0	11.2	2.8
Object Assembly	11.4	3.9	12.0	2.9	11.4	3.0

Bannatyne categories (Bannatyne, 1974) were calculated on the ETS group for comparison with the Cardoni data and are presented in Table III.2. The frequently found pattern of Spatial > Conceptual > Sequential scores for LD students appeared in the Cardoni study. Although the characteristically low score for the Sequential category appeared in the ETS study, the Conceptual score was higher than the Spatial. That result was not surprising in view of the fact that admission to the LD program for students in the ETS study was based in a large part on conceptual skills as demonstrated by the similarities subtest of the WAIS. This selection mechanism operating on the population of LD students applying to the college resulted in a pattern of Conceptual > Spatial > Sequential scores. Nevertheless, scores of the LD groups are more alike than either is to the non-LD group.

Table III.2

Means and Standard Deviations of Bannatyne Categories

	Cardoni et al. Study				ETS LD Study	
	Normal		Learning Disabled		Learning Disabled	
	M	SD	M	SD	M	SD
Spatial	35.7	6.9	35.7	6.0	34.0	6.7
Conceptual	36.6	4.7	35.6	6.1	34.7	5.5
Sequential	35.8	4.2	28.7	5.5	30.9	4.8
Acquired Knowledge	34.4	4.5	29.9	5.5	29.9	5.0

In the next section of this chapter, SAT scores and high school ranks (HSR) will be used to predict the first semester averages of LD and non-LD students. It may be of interest to observe the correlations between WAIS scores and these four variables for the 58 LD students. The data are presented in Table III.3. Most of the WAIS subtests are more

highly correlated with the SAT mathematics score than with the SAT verbal score.

Table III.3

Correlations between WAIS Scores and the SAT, the High School Rank, and the First Semester Average

WAIS Subtests	SAT-V	SAT-M	HSR	FSA
Information	0.25	0.38	0.06	0.28
Comprehension	0.33	0.50	0.18	0.36
Arithmetic	0.35	0.55	0.24	0.03
Similarities	0.27	0.27	0.22	0.18
Vocabulary	0.37	0.43	0.01	0.11
Digit Span	0.21	0.22	-0.10	0.02
Digit Symbol	-0.04	-0.09	-0.05	0.15
Picture Completion	0.12	0.33	-0.07	0.18
Picture Arrangement	0.18	0.39	0.01	0.05
Block Design	0.17	0.34	0.48	0.25
Object Assembly	0.11	0.28	0.19	0.27
Bannatyne Categories				
Spatial	0.17	0.41	0.26	0.30
Conceptual	0.39	0.50	0.20	0.30
Sequential	0.27	0.39	0.06	0.11
Acquired Knowledge	0.40	0.58	0.21	0.17

High school percentile rank (HSR), SAT scores (SAT-V and SAT-M) and first semester college grade point averages (FSA) were obtained from university files for both the learning-disabled students and the rest of the freshman class. These data were used to do a validity study of the SAT for learning-disabled students.

Missing Data

The data base consisted of the records of 179 students, seven of whom were transfer students whose records were not used in the analysis. Further investigation of the remaining 172 revealed that 6 students had values of FSA out of range, and these cases were deleted. Of the remaining 166 cases, 111 were non-LD students and 55 LD students. The 13 deleted

cases revealed no peculiar patterns with regard to HSR and SAT; results of the analysis would in all likelihood have been unaltered if complete data had been available for the 13 cases.

Results

Results on the differential performance between the learning disabled and the rest of the students will be presented first. These will be followed by the presentation of the validity coefficient for each group and then by presentation of the results of regression analysis which address the issues of differential prediction and validity.

Table III.4 presents for both LD and non-LD students the means, standard deviations and t-tests for SAT scores, high school ranks and first semester averages, including the intercorrelations among the variables.

Table III.4

Means, Standard Deviations and t-Statistics for Learning-Disabled and Non-Learning-Disabled Students on SAT-Scores, High School Rank (HSR) and First Semester Average (FSA), Including the Intercorrelation Matrix

Variables	Non-LD (111)	LD (55)	t-Statistic	P
SAT-V	410.63 (78.50)	352.00 (68.46)	4.72	0.000
SAT-M	405.23 (84.13)	381.82 (90.96)	1.64	0.102
HSR	52.40 (24.70)	31.68 (21.27)	5.32	0.000
FSA	2.74 (0.65)	2.30 (0.79)	3.78	0.000

Intercorrelation Matrix

(Non-LD/LD)

Variables	SAT-M	HSR	FSA
SAT-V	0.52/0.68	0.44/0.07	0.34/0.32
SAT-M		0.29/0.17	0.28/0.29
HSR			0.39/0.29

Performance of the LD group is lower than the non-LD on the SAT-V, HSR and FSA; however, they score only slightly lower on the SAT-M. The SAT-V difference is 58 points, but the SAT-M difference is only 28 points which is not significant ($P = 0.10$). The difficulties experienced by the LD in classrooms is quantified by the HSR: on the average, they place at the 32nd percentile compared with the non-LD who place at the 50th percentile. These performances are reflected in the FSA--the LD are approximately one-third of a grade point lower than their classmates. A discriminant analysis was performed using the SAT scores and HSR to see how well these variables could discriminate the learning disabled from the non-learning disabled. (Full results are presented in Table A, Appendix B.) The overall multivariate test confirmed the statistically significant difference between the two groups ($F = 11.7$; 3,169 degrees of freedom; $P < .00005$). The misclassification error of the discriminant was 17%, a relatively small rate which further substantiates the difference between the two groups. Of particular interest is the combined ability of the SAT-V and HSR to cluster the dyslexic points as seen in the scatterplots, Figure B, Appendix B. It is interesting to note that the discriminant function weighted the SAT-V and HSR positively (the higher the discriminant value, the more likely the candidate is learning disabled) while it weighted the SAT-M negatively, indicating that the pattern for learning disability in this study is a high SAT-M and a low SAT-V coupled with a very low high school rank.

Evidence of different patterns of performance for LD versus non-LD students is found in the intercorrelations. The most striking difference is that the correlation between SAT-V and HSR is nearly zero for the

learning disabled, whereas it is 0.44 for the non-learning disabled. Not as dramatic, but in the same direction, is the correlation of HSR with SAT-M--0.17 for the learning disabled versus 0.29 for the others.

Table III.5 presents the validity coefficient of the various prediction systems: the HSGPA alone, the SAT composite (two predictors SAT-V and SAT-M), and the HSGPA together with both SAT scores. The values of the SAT-V and SAT-M validity coefficients for the two groups are not statistically significantly different; however, the HSR validity coefficient is significantly lower for the LD compared to the non-LD. It appears that the HSR is not as useful in predicting performance for the LD as it is for the non-LD at this college. The combined HSR and SAT prediction system has a lower validity coefficient for the learning disabled. The direction of the validity coefficients for the various prediction systems, indicating diminished validity for the learning-disabled population, is interesting but is inconclusive. Validity coefficients for the non-learning disabled of the SAT (.36) and HSR (.39) are lower than those of the general population as reported by Breland (1979). This is not unexpected since HSR is an index of performance that is different from undergraduate grade point average, and FSA, with fewer grades to average, is less stable than the two-semester first year average.

Table III.5

Validity Coefficients of Various Prediction Systems for the Learning-Disabled and Non-Learning-Disabled Students

Prediction System	Validity Coefficient R (R^2) Non-LD (N = 111)	LD (N = 55)	P-Value of Test of Difference between LD and Non-LD Coefficients
1. HSR	0.39 (0.15)	0.29 (0.08)	0.02
2. Composite Test Scores	0.36 (0.13)	0.33 (0.11)	0.14
SAT-V	0.34 (0.12)	0.32 (0.10)	0.16
SAT-M	0.28 (0.08)	0.29 (0.08)	0.24
3. HSR and Tests	0.45 (0.20)	0.42 (0.18)	0.13

The square of the multiple correlation coefficient, R-square, for each prediction system is the percent of variance in FSA explained by that prediction system. The increment in R-square due to adding more predictors is the relative percent explained in the variance as FSA. For the learning disabled, the relative percent of variance in FSA explained by adding SAT to HSR and by adding HSR to SAT is 9% and 6%. The former value is significant ($F = 2.86$, d.f. = 2,51; $P = 0.07$) and the latter value is significant ($F = 3.85$, d.f. = 1,51; $P = 0.06$). For the non-LD, the relative percent of variance in FSA explained by adding SAT to HSR and adding HSR to SAT is 4% and 7%. The relative strengths of the SAT and HSR are opposite to those for the LD: for the SAT, $F = 3.06$, d.f. = 2,107, $P = 0.05$; and for HSR, $F = 8.92$, d.f. = 1,107, $P << .005$. The fact that the increment in R-square due to adding SAT to HSR is greater for the LD than the non-LD is due for the most part to the poorer predictive power of the HSR for the LD students.

All of the above results suggest that although SAT and HSR are valid for predicting college performance, operationally equivalent selection of students based on separate or combined use of the SAT and HSR would yield LD and non-LD groups having internally different patterns of cognitive abilities.

Attention is now turned to the important issue of whether use of prediction systems based on the non-learning-disabled group's data has a negative impact on the learning disabled. This question will be answered in two ways. The first approach will be to estimate and compare two separate prediction formulas, one each for the LD and non-LD groups. These formulas will be based on the combined data across groups, rather than each formula being obtained from separate data sets (see Gulliksen and Wilk, 1950, for a complete description). The reason for this approach is that the parameters are estimated with greater accuracy based on combined data than can be obtained from separate group data.

Until now the analyses have made use of only three variables to predict FSA: HSR, SAT-V, and SAT-M. Separate prediction formulas for the LD and non-LD are estimated by the introduction of a fourth predictor, denoted by HAND, indicating whether or not an individual has a disability. In addition, HAND is multiplied together with each of the other predictors to produce three more variables in order to describe interactions between scores and the handicap. Separate prediction formulas are then obtained for an individual by using his/her value of the predictor HAND. Significant differences in prediction formulas between the LD and non-LD are determined by evaluating the contribution to R-square that the additional predictors (HAND and interactions) make.

Table III.6 shows the R-square values for each of three prediction systems, with and without the variable HAND. Values of the regression weights and related statistics associated with each of the three prediction systems are respectively given in Table C of Appendix B. There were no statistically significant interactions between HAND and the predictors in all three prediction systems; therefore any differences in predictions will be constant for each combination of predictor values. Scatterplots of the data are included as Figure D, Appendix B. Each system is discussed separately.

Table III.6

R-square Value of the Three Prediction Systems

With and Without the Handicap Indicator

	Prediction System	R-Square
1.	HSR	0.18
	HSR and HAND	0.19
2.	Composite Test Score	0.17
	SAT-V	0.11
	SAT-M	0.08
	SAT Composite Test Scores and HAND	0.19
3.	HSR with SAT Composite Test Scores	0.24
	HSR, SAT Composite Test Scores and HAND	0.25

The HSR. The HSR prediction system does not operate equivalently for the LD and non-LD. The contribution to R-square due to the addition of the variable HAND is 0.02, yielding an F-statistic equal to 3.56 with 1 and 163 degrees of freedom (P = 0.03).

By use of the prediction formula based on the non-LD group's data, a learning-disabled student would be predicted to have a higher FSA than his/her own equation would indicate. (See Appendix B, Table C.) A learning-disabled student with HSR equal to 3.2, the LD group mean, would be predicted to obtain an FSA equal to 2.52 based on the non-LD group's data; whereas, the same student would on the average obtain an FSA equal to 2.30. Even though the overall test of different regressions between the two groups is statistically significant for each value of HSR, the difference in predicted FSA, 0.22, is not significant. Since the standard error of estimate is approximately 0.60, the practical differences in predicted FSA would be observed only for outlying values of HSR.

SAT-V and SAT-M. The SAT composite prediction system, like the HSR system, also does not operate equivalently for the LD and non-LD students. In Table III.3, the increment in R-square, 0.03, due to the addition of the variable HAND yields an F-statistic equal to 5.54 with 1 and 162 degrees of freedom (P = 0.01). The net effect of predicting LD students' FSA based on the non-LD group's data is to over predict uniformly over each score combination of SAT-V and SAT-M by 27% of a grade point. However, the standard error of estimate is 0.66, so significant difference in prediction occurs only for combinations of SAT-V and SAT-M far outside the practical range.

The HSR and the SAT. In Table III.3, the prediction system combining the SAT and HSGPA does not operate equivalently for the learning-disabled and non-learning-disabled students. The contribution to R-square due to the addition of the variable HAND is 0.01, yielding a significant F-statistic equal to 1.82 with 1 and 161 degrees of freedom (P = 0.09). The net effect for the LD, resulting from using the HSR and SAT prediction system based on the non-LD data, is to over predict their FSA by 16% of a grade point uniformly over the range of each predictor. The difference in prediction does not appear to be serious since the standard error of estimate is 0.66, indicating that the predicted range of FSA values for given predictor values provides considerable overlapping between the LD and non-LD.

Examination of the residuals revealed no departure from the model, and thus the model adequately described the combined data (Figure D, Appendix B). (Incidentally, the data are appropriately combined since a test of homogeneity of variance yields F = 1.58, 51,107 degrees of freedom; P > .10.) Scattergrams of FSA with each predictor are provided in Figure E, Appendix B.

Discussion

The LD places lower than the non-LD student on the SAT-V and HSR, but performs equally well on the SAT-M. An LD individual's high school performance is not an indicator of performance on the SAT (Appendix B, Figures B.2 and 3). A low high school placement may be incorrectly construed to indicate a failure when, in fact, the LD individual may score reasonably well on the SAT, particularly on the mathematical component. The SAT and HSR predict the FSA of the LD as well as they predict the FSA of the non-LD student. Since the correlational structure of the three predictors, SAT-V, SAT-M and HSR, is different for the two groups, the skills possessed by the

two groups are not the same. This means that selection based on these criteria would yield LD students having verbal, mathematical and other unidentified capabilities in different combinations to non-LD students having otherwise equivalent credentials. The LD student's first semester performance, as suggested by a particular combination of the SAT and HSR, is lower than that attained by a non-LD individual having equivalent credentials. Additional work is necessary to examine the validity of the SAT for learning-disabled students using first-year grade point averages and overall success in college.

Chapter 4

DISCUSSION

These validity studies were undertaken using the hypothesis that all other things being equal, the SAT composite and the HSGPA would predict college GPA equally well for handicapped and nonhandicapped college students.

From the deaf study, we have seen that the SAT scores of deaf students in our study were significantly lower than the scores of their hearing counterparts although both groups had equivalent distributions of HSGPAs. We have seen that verbal SAT scores seem to be related to the age of onset of deafness with its concomitant problems of language development. Although the combined SAT and HSGPA predicted equivalently for both deaf and hearing students at one university, the results for the deaf may be explained on the basis of the combination of over prediction by the HSGPA and under prediction by the SAT composite.

From the learning-disabled study, we have seen that the SAT scores of LD students are lower than those of non-LD students, although only the verbal score was significantly lower. Significant differences for the two groups were also found for high school rank and first semester averages. However, the combined SAT scores and HSR predict the FSA of LD students as well as they predict the FSA of non-LD students.

Several points need to be raised. On the one hand, the studies appear to show that the combination of high school performance and SAT scores has predictive validity for the college performance of both deaf and LD students. On the other hand, performance on the verbal portions of the SAT for both groups and the mathematical section for the deaf is significantly different from the performance of nonhandicapped students.

Special care should be taken that SAT scores for handicapped students are not used in the absence of other information for screening students for admissions or making other admissions decisions. Since this is already the policy of the Admissions Testing Program of the College Board, the statement is not new. But to the extent that cutoff scores for the SAT exist at any university, that mechanism may have a greater impact on deaf and learning-disabled students than on the nonhandicapped, if the results of these studies can be generalized to other institutions.

To return to our initial hypothesis that, all other things being equal, the SAT composite and a measure of high school performance would predict college performance equally well for handicapped and nonhandicapped students, all other things were not equal. Both institutions which participated in the validity studies had outstanding programs of support for students with a specific disability who, nevertheless, were mainstreamed within the institution. There are few, if any, locations in regular college and university settings where support services are more thoroughly provided and special needs met. In addition, no attempt was made in the current studies to correct for the amount of course work completed for the college performance measure of the FSA or FYGPA. It is quite likely that deaf and LD students took at least one less course per semester than nonhandicapped students. Such advice was often given. Without the kind of support services provided by the institutions in our study, the validity data in this report might have been quite different. Further work is necessary if we are to understand fully the implications of standardized tests for disabled persons.

Next Steps

Future research into the validity of tests for handicapped populations could include similar studies as those presented here, utilize Modern Bayesian statistical approaches, or study item-level responses. These options would not have to be performed individually, but could be performed in various combinations. These kinds of studies are needed to insure that equivalent information is available on the handicapped as well as the nonhandicapped.

The areas of research just mentioned touch on some of the many problems connected with validity studies on handicapped populations. One of these problems is finding students with particular kinds of handicaps having matriculated in colleges that enroll both the handicapped and nonhandicapped. Another problem is pooling these data over schools to determine adequately the predictive validity for the handicapped. Still another problem involves the proper identification of a particular handicapped group of students and the subsequent categorization so as to contrast them intelligently with nonhandicapped groups. Many problems are not yet known and waiting to surface once more studies are attempted. Consequently, research programs need to remain flexible by allowing the application of new research tools as they become available and to develop a system to accumulate data on the handicapped whose importance is not immediately apparent.

For the present, some specific suggestions involving available data bases and promising research tools are presented.

Similar Studies

The predictive validity study on the deaf presented here uses the first year average as a criterion. Available on these deaf students is

the grade point average for each year, and thus a validity study involving an overall criterion of college success is feasible.

The predictive study on the learning disabled used the first semester average. Other criteria could become available as these students progress through college. The first year grade point average may already be available.

Similar studies involving large groups of disabled persons, such as the blind or physically handicapped, could be arranged by systematically contacting a sufficient number of colleges.

Bayesian Statistics

In the event that only small numbers of any particular handicap at one college are available but over a large number of colleges, Bayesian statistical methods could be applicable. One such study involving small groups of minorities in graduate business schools is Braun and Jones (1981).

Item-Level Analyses

Construct validity and differential performance at the item level are feasible on data sets having a sufficient number of suitably identified handicaps. One exemplary study involving the deaf is Allen (1981), and another involving the learning disabled is Bannatyne (1974).

References

Allen, T. E. Test research variation between hearing impaired and hearing students. Research Report. Washington, D.C.: Gallaudet College Research Institute.

Bannatyne, A. Diagnosis: A note on recategorization of the WISC scaled scores. Journal of Learning Disabilities, 1974, 7, 272-273.

Braun, H. I., & Jones, D. H. Graduate Management Admission Test predictive bias. To appear as GMAC Research Report. Princeton, N.J.: Educational Testing Service.

Breland, H. M. Population Validity and College Entrance Measures. Research Monograph No. 8. New York: The College Board, 1979.

Cardoni, B. K., O'Donnell, J. P., Ramaniah, N. V., Kurtz, J., & Rosenshein, K. Journal of Learning Disabilities, Vol. 14, No. 7, August/September 1981.

Federal Register. Department of Health, Education, and Welfare, Office of the Secretary. Non-Discrimination on Basis of Handicap. Title 45-Public Welfare, May 4, 1977. Part IV, pp. 22676-22702.

Gulliksen, H., & Wilks, S. S. Regression test for several samples. Psychometrica, 1950, 15, 91-114.

Morrison, D. E. Multivariate Statistical Methods. New York: McGraw-Hill 1976.

Panel on Testing of Handicapped People, National Academy of Sciences, Washington, March 14-15, 1980.

Rubin, D. B. Inference and Missing Data. Biometrika, 1976, 63, 581-92.

Searle, S. R. Linear models. New York: Wiley, 1973.

Snedecor, G. W., and Cochran, W. G. Statistical methods (6th ed.). Ames, Iowa: The Iowa State University Press, 1973.

The Program of Assistance in Learning, The Robert Smith Memorial Invitational Conference, Milton, Mass., May 8, 1981.

Wechsler, D. Manual: Wechsler Adult Intelligence Scale. New York: Psychological Corp., 1955.

APPENDIX A

Tables and Figures, Deaf Validity Study

Table A

Discriminant Function and Summary Statistics

Variable	Weight (x1/10,000)	Means (S.D.) Hearing (N = 264)	Deaf (N = 109)	Difference	Approximate Probability of Misclassification Using This Variable Alone %
SAT-V	10.1190	455.61 (92.56)	358.26 (112.88)	97.35 (98.91)	30.8
SAT-M	1.9973	480.80 (100.33)	413.94 (100.96)	66.86 (102.65)	40.1
HSGPA	-13.9350	3.22 (0.59)	3.18 (0.62)	0.04 (0.60)	99.9

A1

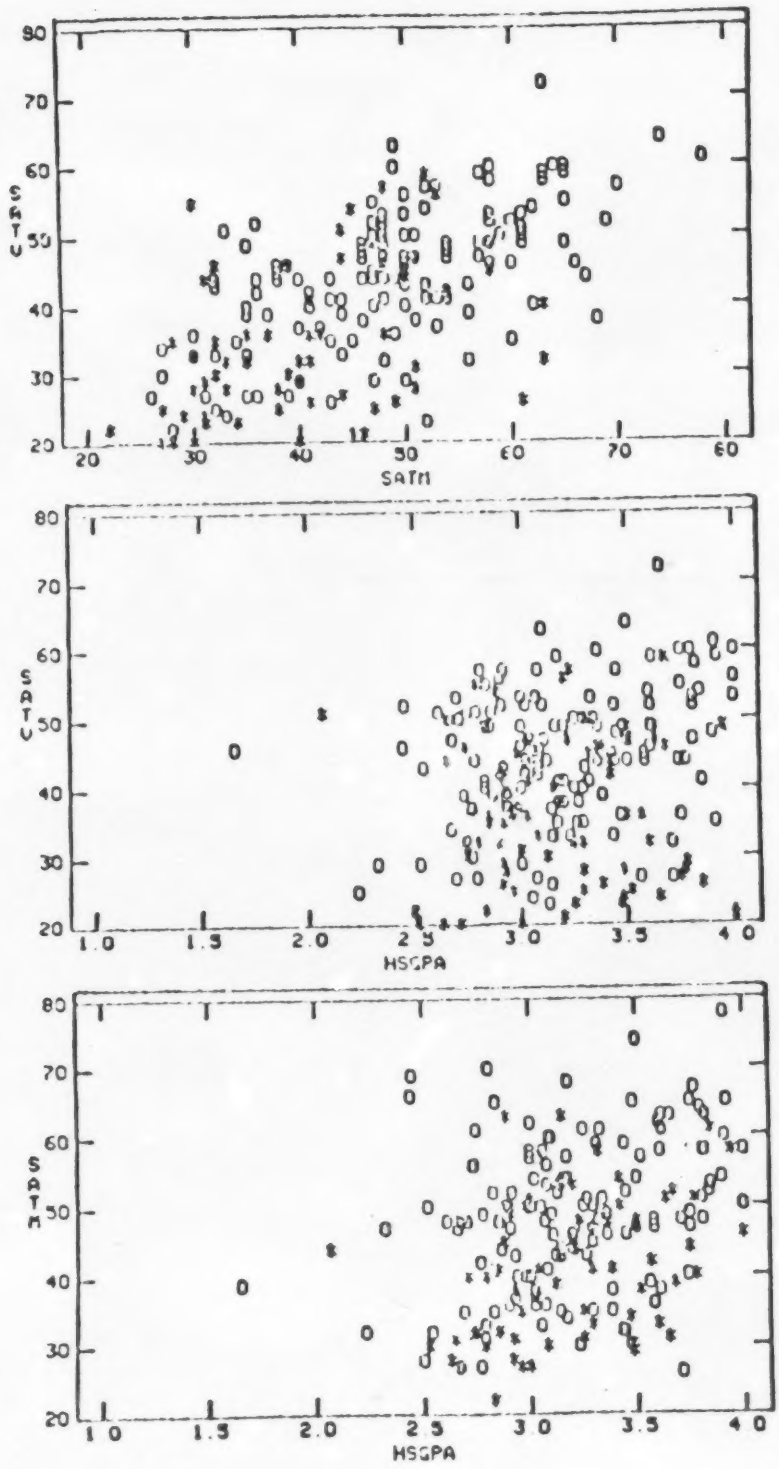


Figure B: Scatterplots of the predictor variables. Scores for the deaf and hearing students are represented by * and 0 respectively.

Table C

Values for Each of Three Prediction Systems, Dependent Variable Is FYA

Prediction System	Standard Reg. Weight	Regression Weights	Standard Error of Wt.	t-statistics	D.F.	Contribution to R-Sq.	Multiple Corr.	Standard Err. Est.
1. HSGPA								
Constant		-0.8417	0.3621	-2.3243	196		0.6134	0.5549
Independent Variables								
HSGPA	0.6433	1.0944	0.1123	9.7470		-0.3024		
HGPA1HAND	-0.6091	-0.2860	0.2163	-1.3223		-0.0056		
HANDICAP	0.4881	0.7408	0.6972	1.0626		-0.0036		
2. SAT								
Constant		1.2064	0.2845	4.2409	194		0.4130	0.6432
Independent Variables								
SAT-MATH	0.1497	0.0094	0.0064	1.4783		-0.0093		
SAT-VERB	0.3529	0.0221	0.0069	3.1840		-0.0433		
HANDICAP	0.1892	0.2872	0.4748	0.6048		-0.0016		
HAND/S-M	0.4589	0.0163	0.0112	1.4606		-0.0091		
HAND/S-V	-0.5838	-0.0239	0.0109	-2.1936		-0.0206		
3. SAT and HSGPA								
Constant		-1.2176	0.3747	-3.2492	192		0.6481	0.540
Independent Variables								
SAT-MATH	0.0565	0.0036	0.0054	0.6585		-0.0013		
SAT-VERB	0.2087	0.0131	0.0059	2.2032		-0.0147		
HSGPA	0.5727	0.9743	0.1160	8.4009		-0.2132		
HANDICAP	0.5554	0.8430	0.7297	1.1554		-0.0040		
HAND/S-M	0.2967	0.0106	0.0100	1.0557		-0.0034		
HAND/S-V	-0.2474	-0.0101	0.0093	-1.0865		-0.0036		
HAND/HSGPA	-0.6238	-0.2929	0.2281	-1.2845		-0.0050		

A3

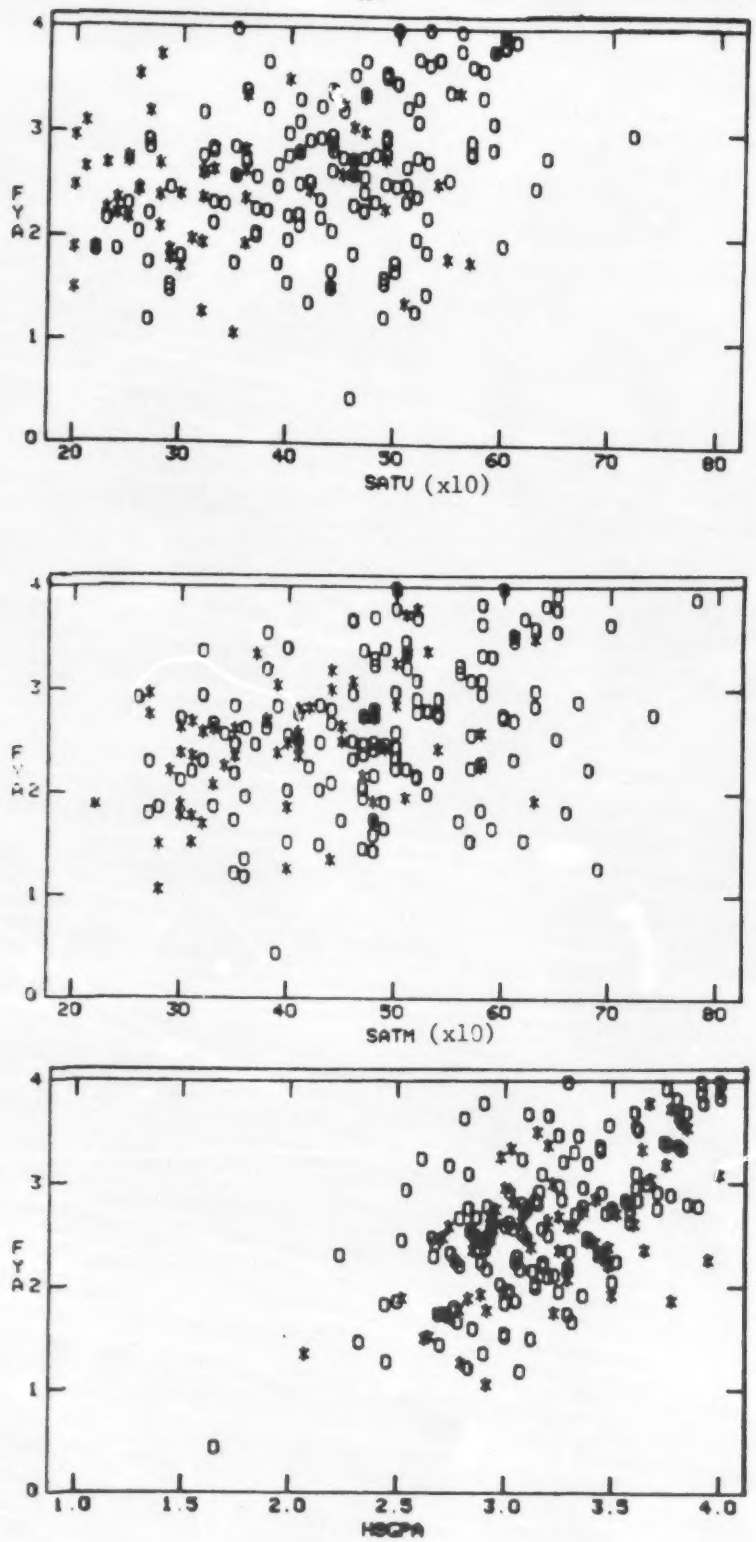
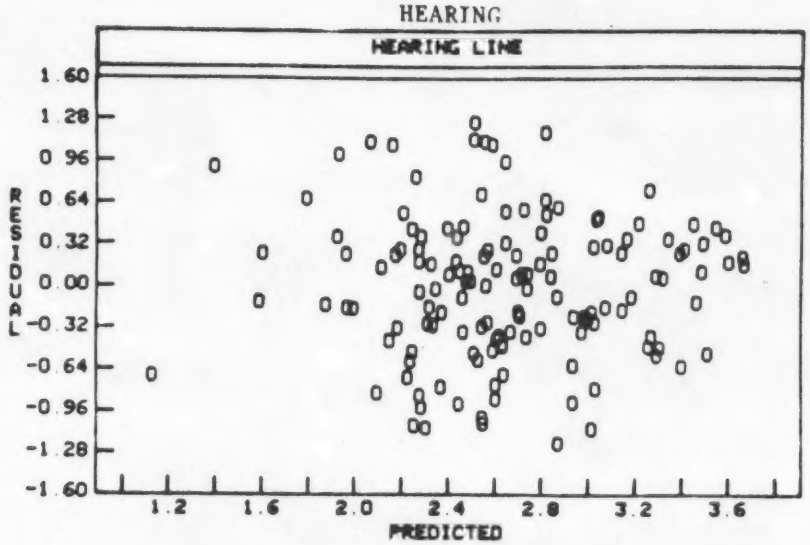
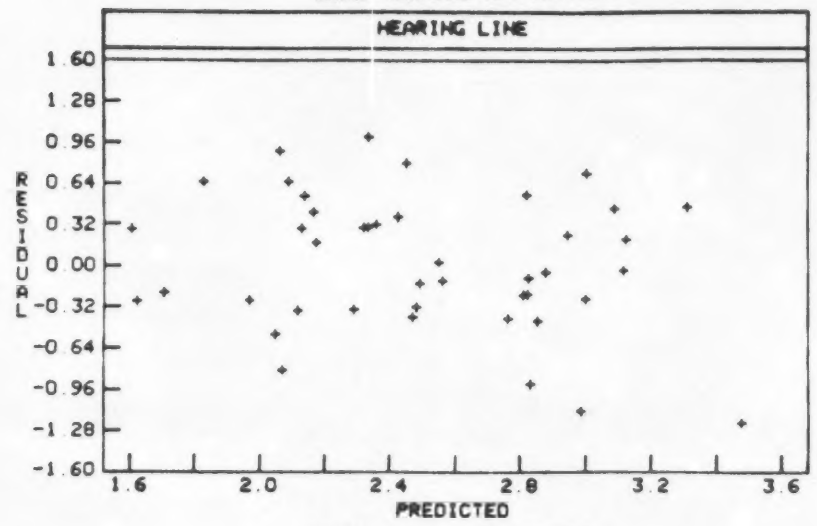


Figure D: Scatterplots of FYA versus the predictor variables SAT-V, SAT-M and HSGPA reveal that SAT-V underpredicts FYA for the deaf as compared to the hearing, but there is no apparent difference in the SAT-M and HSGPA. Deaf students are denoted by * and hearing by 0.

HEARING



DEAF BEFORE 18 MONTHS



DEAF AFTER 18 MONTHS

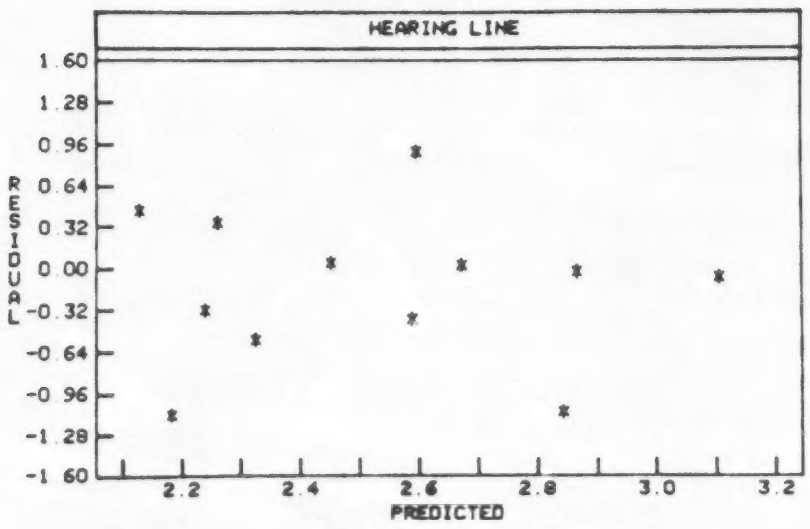


Figure E: Scatterplots of residuals

(Observed-Predicted)

APPENDIX B

Tables and Figures, Learning Disabled Validity Study

STATISTICAL ANALYSES FOR THE DATA ON DYSLEXIC STUDENTS FROM
DISCRIMINATE ANALYSES

MODEL 1.3

Table A

THE DEPENDENT VARIABLE IS Y , THE MULTIPLE CORRELATION IS 0.4157. THE STANDARD ERROR OF ESTIMATE= 0.0113

	SUM OF SQUARES	PROPORTION OF SQUARES	N.D.F.	MEAN SQUARE	F RATIO	PROBABILITY OF LARGER F
TOTAL ABOUT ORIGIN	0.0259		173.			
UNDER NULL HYPOTHESIS	0.0259	1.0000	172.			
DUE TO HYPOTHESIS	0.0045	0.1728	3.	0.0015	11.7654	0.0000
ERROR	0.0215	0.8272	169.	0.0001		

	STANDARD REG. WEIGHT	REGRESSION WEIGHTS	STANDARD ERROR OF WT.	T STATISTICS WITH 169.D.F.	CONTRIBUTION TO R-SQ.	MEASURE OF COLLINEARITY
CONCOMITANT VARIABLES						
PSUEDO		-0.0158	0.0047	-3.3417		0.9674
INDEPENDENT VARIABLES						
SAT VERB	0.2377	0.0000	0.0000	2.6590	-0.0346	0.3874
SAT MATH	-0.0857	-0.0000	0.0000	-1.0098	-0.0050	0.3197
HS-RANK	0.2972	0.0001	0.0000	3.8803	-0.0737	0.1657

REGRESSION WEIGHTS

PSUEDO	-0.01584634
SAT VERB	0.00003673
SAT MATH	-0.00001229
HS-RANK	0.00010000

STANDARD ERROR OF WT.

PSUEDO	0.00472809
SAT VERB	0.00001377
SAT MATH	0.00001214
HS-RANK	0.0

MEAN SQUARE

Y	0.00012621

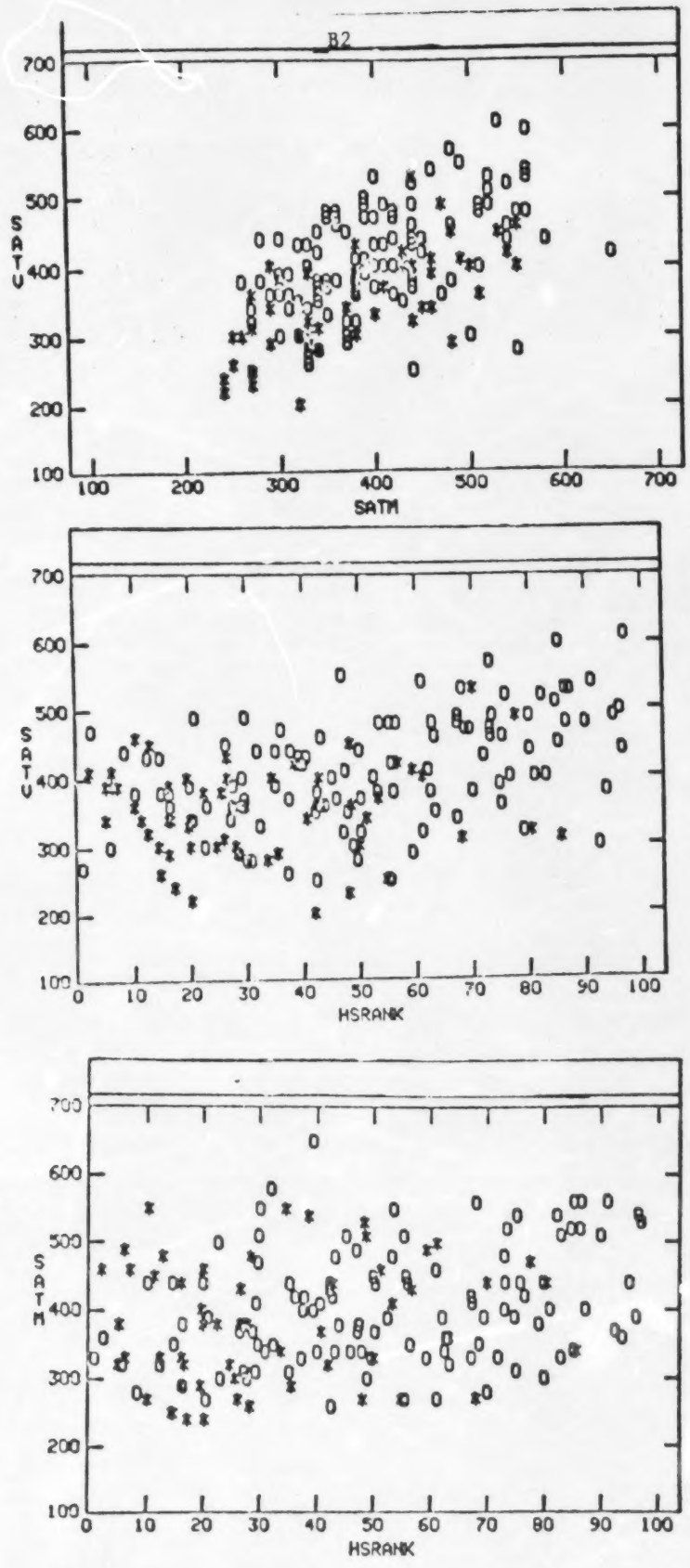


Figure B: Scatterplots of the predictor variables. Scores for the LD and nonLD students are represented by * and 0 respectively.

Table C

Values for Each of Three Prediction Systems, Dependent Variable Is FSA

Prediction System	Standard Reg. Weight	Regression Weights	Standard Error of Wt.	t-statistics	D.F.	Contribution to R-Sq.	Multiple Corr.	Standard Err. Est.
1. HSR								
Constant		2.1956	0.1296	16.9468	163		0.4392	0.6559
Independent Variables								
HSR	0.3635	0.0103	0.0022	4.7697		-0.1127		
HANDICAP	-0.1438	-0.2210	0.1171	-1.8869		-0.0176		
2. SAT								
Constant		1.3236	0.3035	2.7786	162		0.4394	0.6578
Independent Variables								
SAT-VERB	0.2541	0.0023	0.0008	2.7786		-0.0385		
SAT-MATH	0.1383	0.0012	0.0007	1.5985		-0.0127		
HANDICAP	-0.1778	-0.2732	0.1161	-2.3533		-0.0276		
3. SAT and HSR								
Constant		1.2809	0.2941	4.3557	161		0.4983	0.6369
Independent Variables								
SAT-VERB	0.1766	0.0016	0.0008	1.9328		-0.0174		
SAT-MATH	0.1167	0.0010	0.0007	1.3891		-0.0090		
HSR	0.2711	0.0077	0.0022	3.4376		-0.0552		
HANDICAP	-0.1033	-0.1587	0.1172	-1.3546		-0.0086		

B3

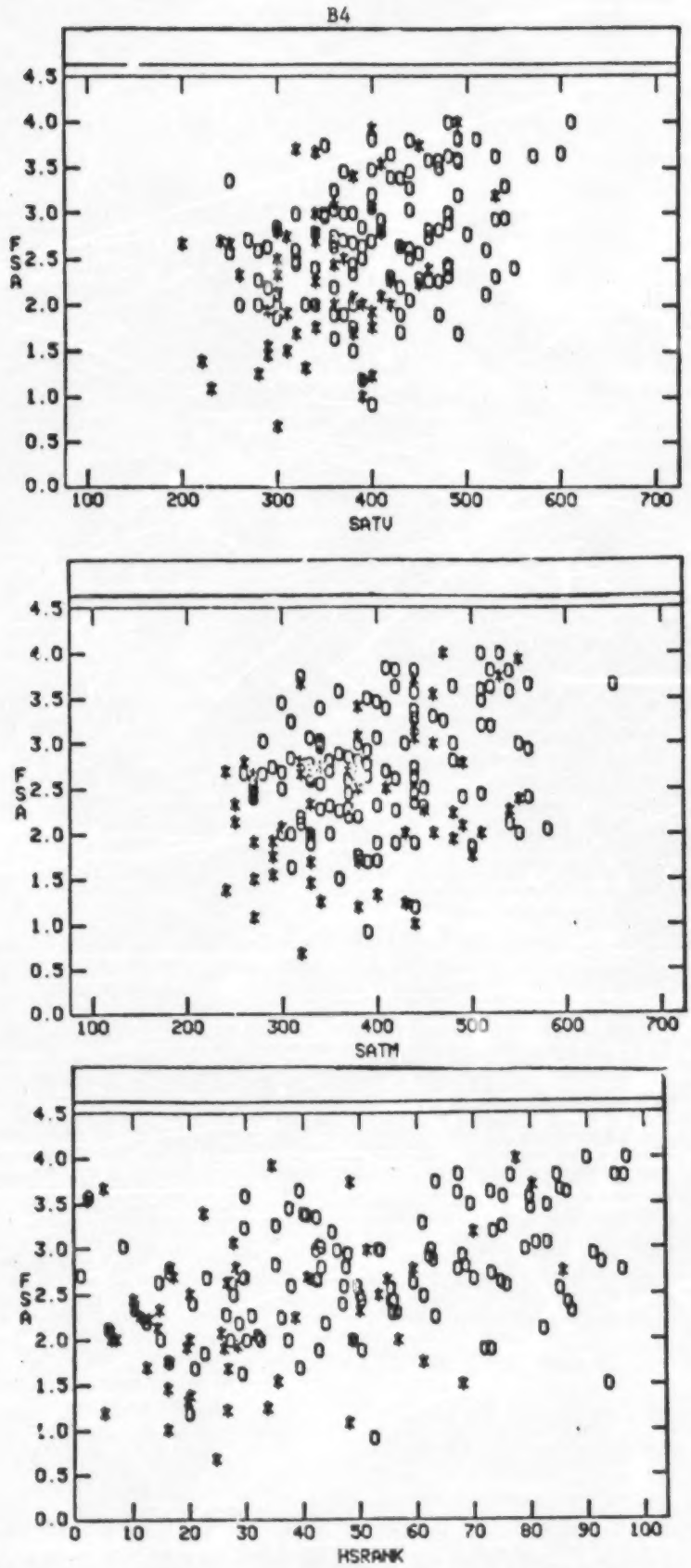


Figure D: Scatterplots of FSA versus the predictor variables SAT-V, SAT-M and HSR. LD and nonLD students are denoted by * and 0 respectively.

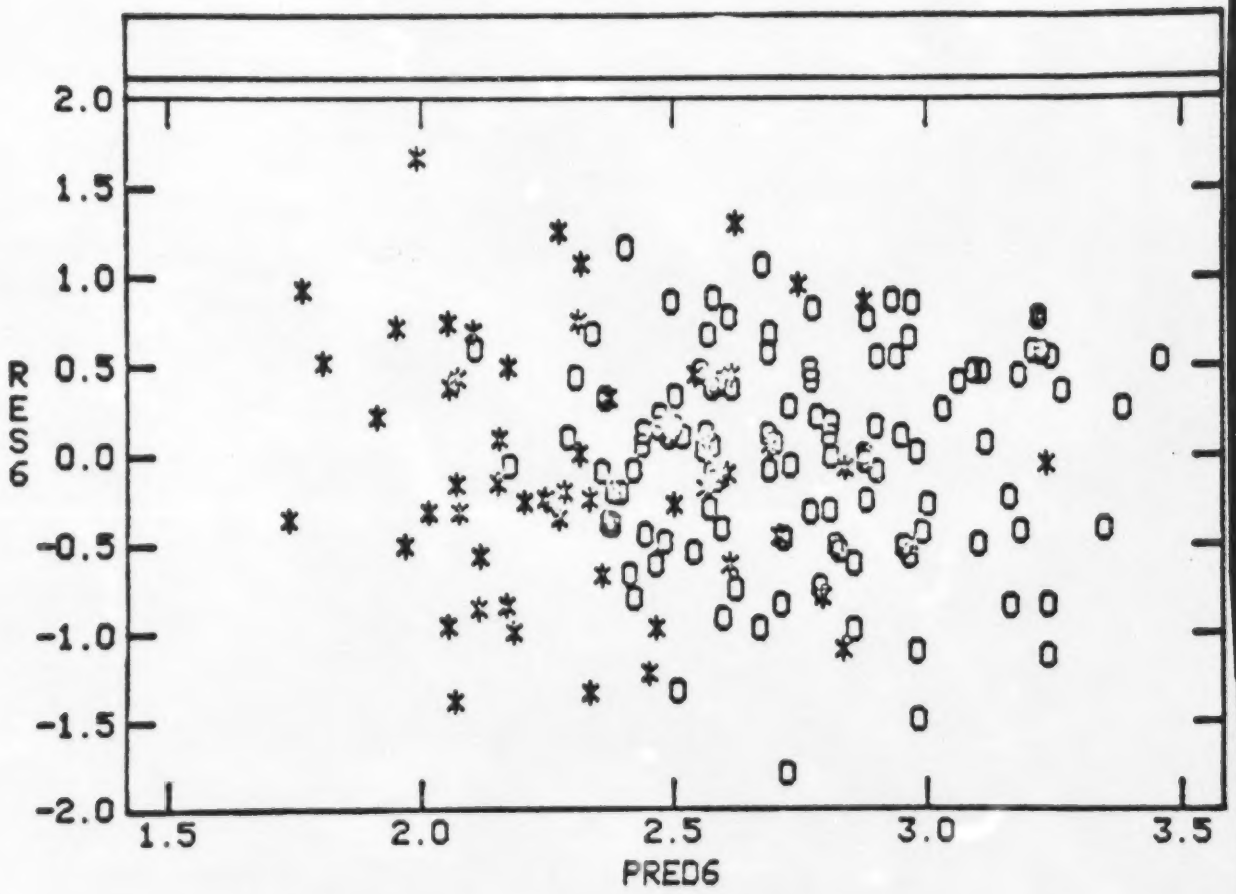


Figure E: Residual versus predicted FSA plot using HSR and SAT prediction system. LD and non-LD students are denoted by * and 0 respectively.